MAJOR ARTICLE

Molecular Epidemiology Reveals Long-Term Changes in HIV Type 1 Subtype B Transmission in Switzerland

Roger D. Kouyos,[1,a] Viktor von Wyl,[1,a] Sabine Yerly,[4] Jürg Böni,[2] Patrick Taffé,[6] Cyril Shah,[2] Philippe Bürgisser,[7] Thomas Klimkait,[8] Rainer Weber,[1] Bernard Hirschel,[5] Matthias Cavassini,[7] Hansjakob Furrer,[10] Manuel Battegay,[9] Pietro L. Vernazza,[11] Enos Bernasconi,[12] Martin Rickenbach,[6] Bruno Ledergerber,[1] Sebastian Bonhoeffer,[3] Huldrych F. Günthard,[1] and the Swiss HIV Cohort Study

[1]Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, [2]Swiss National Center for Retroviruses, University of Zurich, and [3]Institute of Integrative Biology, Eidgenössische Technische Hochschule Zurich, Zurich, [4]Central Laboratory of Virology and [5]Division of Infectious Diseases, Geneva University Hospital, Geneva, [6]Swiss HIV Cohort Study Data Center, [7]Division of Immunology, University Hospital Lausanne, Lausanne, [8]Institute for Medical Microbiology, University of Basel, [9]Division of Infectious Diseases and Hospital Epidemiology, University Hospital Basel, Basel, [10]Division of Infectious Diseases, University Hospital Berne, Berne, [11]Division of Infectious Diseases, Cantonal Hospital St. Gallen, St. Gallen, [12]Division of Infectious Diseases, Regional Hospital Lugano, Lugano, Switzerland

Background. Sequence data from resistance testing offer unique opportunities to characterize the structure of human immunodeficiency virus (HIV) infection epidemics.

Methods. We analyzed a representative set of HIV type 1 (HIV-1) subtype B pol sequences from 5700 patients enrolled in the Swiss HIV Cohort Study. We pooled these sequences with the same number of sequences from foreign epidemics, inferred a phylogeny, and identified Swiss transmission clusters as clades having a minimal size of 10 and containing $\geqslant$80% Swiss sequences.

Results. More than one-half of Swiss patients were included within 60 transmission clusters. Most transmission clusters were significantly dominated by specific transmission routes, which were used to identify the following patient groups: men having sex with men (MSM) (38 transmission clusters; average cluster size, 29 patients) or patients acquiring HIV through heterosexual contact (HETs) and injection drug users (IDUs) (12 transmission clusters; average cluster size, 144 patients). Interestingly, there were no transmission clusters dominated by sequences from HETs only. Although 44% of all HETs who were infected between 1983 and 1986 clustered with injection drug users, this percentage decreased to 18% for 2003–2006 ($P < .001$), indicating a diminishing role of injection drug users in transmission among HETs over time.

Conclusions. Our analysis suggests (1) the absence of a self-sustaining epidemic of HIV-1 subtype B in HETs in Switzerland and (2) a temporally decreasing clustering of HIV infections in HETs and IDUs.

The fast evolution of human immunodeficiency virus (HIV) enables the inference of important epidemiologic patterns from sequence data [1]. Phylogenetic methods can address questions that are very difficult to investigate otherwise, such as investigations of the early phase of the epidemics, for which few direct data are available [2–5]. Furthermore, molecular methods have been applied in identifying transmission events [6] and characterizing large transmission chains [4, 7–11].

Received 25 June 2009; accepted 16 November 2009; electronically published 12 April 2010.

Reprints or correspondence: Dr Roger D. Kouyos, Theoretical Biology, Universitaetsstrasse 16, ETH Zentrum, CHN H 75.1, CH-8092 Zurich, Switzerland (roger.kouyos@env.ethz.ch); Dr Huldrych F. Günthard, University Hospital Zurich, Div of Infectious Diseases and Hospital Epidemiology, Raemistrasse 100, 8091 Zurich, Switzerland (huldrych.guenthard@usz.ch).

The Journal of Infectious Diseases 2010; 201(10):1488–1497

0022-1899/2010/20110-0008$15.00
DOI: 10.1086/651951

Potential conflicts of interest: H.F.G. has been an adviser and/or consultant for GlaxoSmithKline, Abbott, Novartis, Boehringer-Ingelheim, Roche, Tibotec, and Bristol-Myers Squibb and has received unrestricted research and educational grants from Roche, Abbott, Bristol-Myers Squibb, GlaxoSmithKline, Tibotec, and Merck Sharp & Dohme. S.Y. has participated on the advisory boards of Bristol-Myers Squibb and Tibotec and has received travel grants from GlaxoSmithKline and Merck Sharp & Dohme.

Presented in part: 16th Conference on Retroviruses and Opportunistic Infections, Montréal, Canada, 8–11 February 2009 (poster 286); 16th International Workshop on HIV Dynamics & Evolution, Oxford, United Kingdom, 4–7 April 2009 (abstract 17).

Financial support: See the Acknowledgments.

[a] Both authors contributed equally to this article.

Switzerland belongs to the Western European countries that have been most strongly affected by HIV type 1 (HIV-1), as indicated by the sustained higher prevalence of HIV-1 noted in Switzerland, compared with its neighboring countries of Germany, France, and Italy [12]. The structure of the Swiss epidemic is typical of such epidemics seen in Western Europe. The predominant transmission routes were used to identify the following patient groups: injection drug users (IDUs), men who have sex with men (MSM), and individuals acquiring HIV through heterosexual intercourse (HETs). Although HIV-1 transmission preferentially occurs among individuals with the same risk factor, interactions between transmission groups (eg, prostitution among IDUs or bisexuality) might play an important role as well. An additional and increasingly more important layer of complexity in the study of disease dynamics emerges from infections acquired abroad, either through travel to foreign countries or immigration [13]; this complexity is also evidenced by the increasing proportion of non-B subtypes noted among recently infected patients over the past years [14]. Disentangling these factors is crucial for understanding the driving forces of the epidemic and, thereby, for developing targeted prevention of HIV transmission.

Genotypic drug resistance testing has been routinely performed in Switzerland since 2000. This practice has led to a wealth of sequence data. Moreover, the data have been considerably complemented by retrospective sequencing of stored blood samples. The sampling times for these retrospectively sequenced samples go back to 1995, and the years of enrollment of the patients covered reach back to 1984. Sequence data have been linked to the Swiss HIV Cohort Study (SHCS), which collects clinical data from a representative fraction (49%) of the HIV-infected individuals living in Switzerland [15]. The unique features of the SHCS are the large proportion of women and the good representation of all transmission groups (IDUs, MSM, and HETs). The SHCS drug resistance database contains sequence data for 56% of the 15,000 patients who ever enrolled in the SHCS and thus provides a representative picture of the HIV epidemic in Switzerland.

Using phylogenetic clustering methods, we sought to identify and characterize chains of transmission of HIV-1 subtype B within Switzerland. In particular, we set out to analyze mixing patterns between transmission groups and the associated temporal trends.

MATERIALS AND METHODS

Sequences and patients. The SHCS is a nationwide, prospective, clinic-based cohort study with continuous enrollment and semiannual study visits [16]. The SHCS has been approved by the ethics committees of all participating institutions, and written, informed consent has been obtained from participants. The SHCS drug resistance database contains all HIV resistance tests performed by the 4 laboratories engaged in HIV resistance testing in Switzerland and stored in a central database developed and hosted by SmartGene (Integrated Database Network System version 3.5.0; SmartGene) [17]. Resistance data stem from routine clinical testing (60% of tests) and from tests retrospectively performed using frozen repository plasma samples (40% of tests). Retrospective sequencing was systematically performed using the oldest available plasma sample from each patient. All laboratories are performing population-based sequencing of the full protease gene and, at a minimum, codons 28–225 of the reverse-transcriptase gene, by use of commercial assays (Viroseq, version 1 [PE Biosystems]; Virsoseq, version 2 [Abbott]; and vircoTYPE HIV-1 Assay [Virco Lab]) and in-house methods [18].

The SHCS drug resistance database contains 11,841 sequences from 8572 patients with at least partial protease and reverse-transcriptase sequences available. The sequence lengths range from 900 to 1497 bp. The predominant subtype of the Swiss HIV-1 epidemic is subtype B [14, 19], and therefore we restricted our sample to subtype B viruses ($n = 9157$ sequences). Only the oldest sequence available from each patient was included in the data set, leaving 5700 sequences. To identify clusters of Swiss sequences, we added the same number of randomly selected non-Swiss subtype B sequences from the Los Alamos Sequence database [20]. These sequences stem mainly from the United States (proportion, 34%), Italy (16%), Spain (5%), Canada (5%), Brazil (5%), Great Britain (4%), Argentina (4%), Germany (4%), France (3%), Hungary (3%), and 47 other countries (18%). To avoid distortion of our analysis resulting from convergent evolution driven by antiretroviral therapy, we removed all major amino acid positions that are associated with antiretroviral drug resistance, in accordance with the International AIDS Society (IAS)–USA guidelines [21] (positions 30, 32, 33, 46–48, 50, 54, 76, 82, 84, 88, and 90 in the protease and 41, 62, 65, 67, 69, 70, 74, 75, 77, 100, 103, 106, 108, 115, 116, 151, 181, 184, 188, 190, 210, 215, 219, 225, and 236 in the reverse transcriptase).

The sequences analyzed in the present study constitute a data set that is very representative of an entire country. In recent study years, the sampling density has reached almost 80% of all patients with an AIDS diagnosis. Such data would, in principle, allow for the reconstruction of entire transmission networks and could thereby endanger the privacy of the patients. This is especially problematic because HIV-1 sequences frequently have been used in court cases. From a scientific point of view, the consequences of open and uncontrolled access to such densely sampled sequences could jeopardize the future publication (and, thus, the investigation) of similarly complete data sets and could thereby be counterproductive even from an "open-access" perspective. For these reasons, we decided to make only a random subset of 10% of the sequences accessible

via GenBank (accession numbers, GU344102–GU344671). We would, however, like to point out that all data in the SHCS can be used for well-defined projects that are in accordance with the guidelines of the SHCS, if a corresponding project proposal is approved by the SHCS scientific board.

Demographic data (on transmission groups and geographical origin in Switzerland) and administrative data (on the year that sampling was performed, the year of the enrollment of the patient in the cohort, and the year that the first positive HIV test result occurred) were obtained from the SHCS database. The year of infection was estimated as described elsewhere [22]. In brief, dates were obtained from a back-calculation model that incorporates the dates at which the first positive or last negative HIV test results and CD4 cell counts were obtained as predictor variables. These dates could be calculated for 4563 (80%) of 5700 patients. Patients were categorized into the following groups on the basis of transmission routes: IDUs (1773 sequences), MSMs (2547 sequences), and HETs (1390 sequences). Other transmission modes (eg, perinatal and transfusion risks) were not considered. Years of sampling ranged from 1989 to 2007 (median sampling year [interquartile range {IQR}], 2000 [1997 to 2003]), whereas the year of enrollment into the SHCS ranged from 1984 to 2008 (median year of enrollment [IQR], 1997 [1992 to 2001]). The estimated year of infection ranged from 1981 to 2007 (median year of infection [IQR], 1992 [1988 to 1997]).

Phylogenetic methods. We inferred a maximum-likelihood tree, using the GTRCAT [23] method (which is a computational approximation of the GTR model with Γ-distributed rate heterogeneity) implemented in RAxML [24]. We verified the results inferred from this tree on several trees derived on the basis of alternative methods (see Discussion). Transmission clusters were inferred from this tree in a way similar to that described by Hue et al [4]: "Swiss transmission clusters" were defined as clades that consisted of $\geqslant$80% sequences from Switzerland and that contained $\geqslant$10 sequences. The minimal size was chosen to exclude clusters that would also occur if Swiss and foreign patients were randomly distributed on the phylogeny: with increasing cluster size, clusters with 80% sequences from Switzerland become less likely by chance; for a size of 10, the probability of finding such a cluster by chance is 0.05 (under a binomial model).

RESULTS

A maximum-likelihood tree was inferred from the 11,400 sequences of our data set (5700 sequences from the SHCS plus the same number of sequences from the Los Alamos HIV database). In this tree, we identify Swiss transmission chains (or subepidemics) as clades (or subtrees) that have a minimal size of 10 and for which $\geqslant$80% of the sequences stem from Switzerland (see Materials and Methods and [4]). The intuition behind this definition is that if a group of Swiss patients stems from a Swiss transmission chain, then the sequences from these patients should be more closely related to each other than to sequences from foreign epidemics. Hence, these sequences should form a clade in a phylogeny, even if this phylogeny contains sequences from foreign epidemics. According to this reasoning, Swiss transmission chains can be identified as clades of Swiss sequences in a phylogeny that contains sequences from both Switzerland and abroad. There are, however, at least 2 reasons why subepidemics in Switzerland do not have to lead to clades that are composed uniquely of Swiss sequences. First, individual patients can either emigrate or infect patients outside of Switzerland. These non-Swiss patients will belong to the same clades as the Swiss patients, even if the latter patients form a largely domestic transmission chain. Second, it is well known that HIV trees are subject to a large topological uncertainty (see Discussion); therefore, occasional misattributions of individual patients to wrong clades are likely to occur. Intuitively, the inference of transmission chains from clades should therefore be robust against these 2 processes. One possibility to achieve this is to identify Swiss transmission chains not as clades that are composed uniquely of Swiss sequences but, rather, as clades for which the fraction of Swiss sequences exceeds some (large) threshold. The concrete choice of such a threshold is, to some extent, ad hoc (with the restriction that it should considerably exceed the average fraction of Swiss sequences in the entire phylogeny). For the present study, we have chosen a threshold of 80%, but we have found qualitatively identical results for different values of this threshold (results not shown). In summary, we identify transmission chains as clades dominated by Swiss sequences. In the remainder of the text, we will refer to these as (Swiss) transmission clusters.

Sequences cluster according to transmission group. Figure 1 highlights the distribution of the Swiss sequences in the phylogeny. It indicates the presence of 2 types of transmission clusters: the "IDU/HET transmission clusters," which mainly consist of patients who have acquired HIV through heterosexual contact or injection drug use, and the "MSM transmission clusters," which mainly consist of patients for whom homosexual intercourse between men is the primary route of transmission.

Overall, we found 60 Swiss transmission clusters containing 3120 (55%) of all SHCS sequences (Table 1 characterizes the 20 largest clusters, with a minimum cluster size of 30 individuals). Of these 60 transmission clusters, 50 contain a significant (according to a binomial test) excess of either MSM (38 transmission clusters; 1100 patients) or IDUs and HETs (12 transmission clusters; 1725 patients). Thus, IDUs and HETs segregated into fewer but larger transmission clusters (mean size, 144 individuals; 82% of patients in IDU/HET transmission clusters belonged to the 3 largest transmission clusters), com-

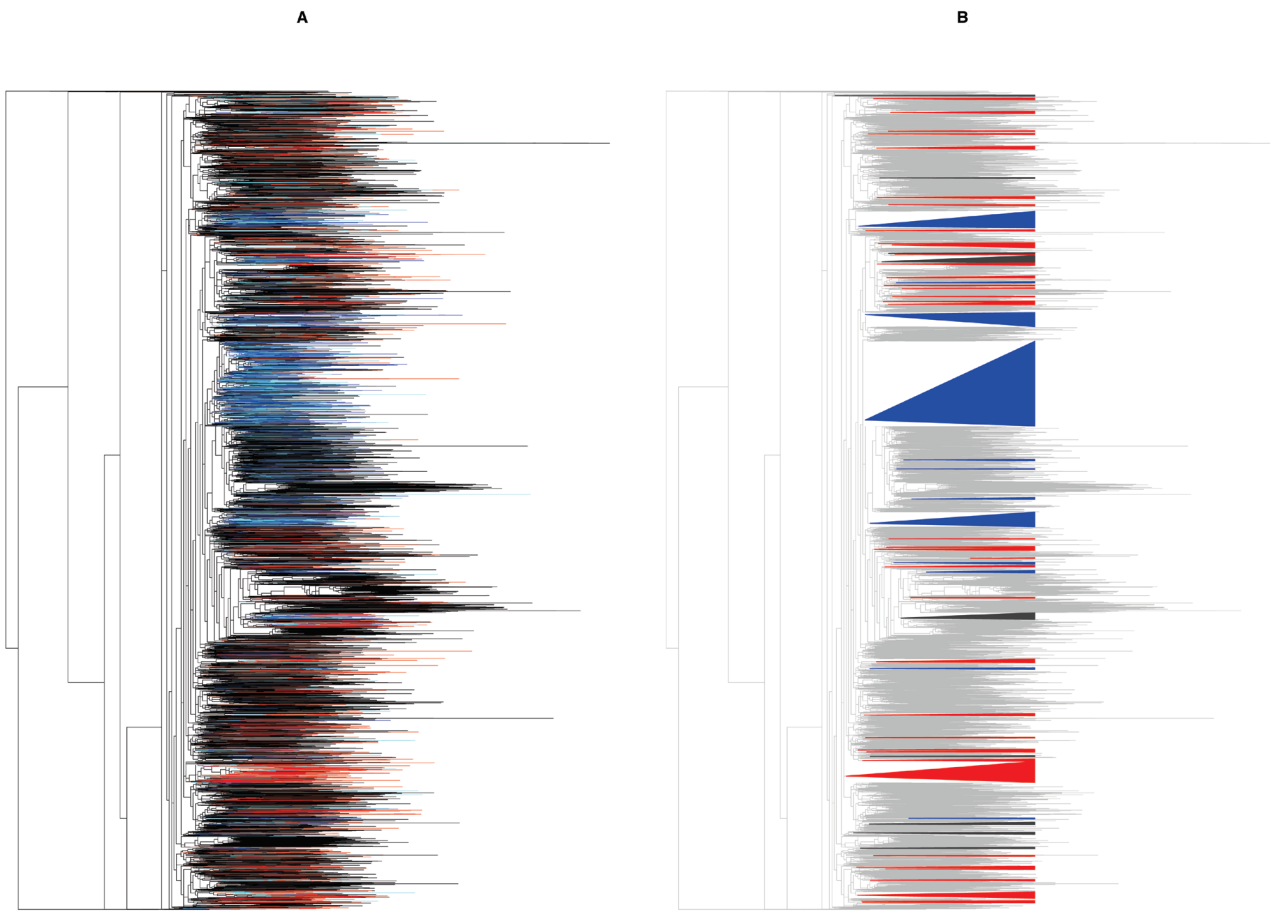

Figure 1. *A,* Maximum-likelihood tree of the entire sequence data set (30 subtype C sequences from the Swiss HIV Cohort Study [SHCS] data set were used as an outgroup). Colors denote the origin and risk group: black denotes non-Swiss origin, cyan denotes Swiss origin and men who have sex with men (MSM), red denotes Swiss origin and patients acquiring human immunodeficiency virus via heterosexual contact (HET), blue denotes Swiss origin and injection drug user (IDU). *B,* Same tree as shown in panel *A* but in which Swiss transmission clusters are denoted by triangles. Red denotes significant dominance of MSMs, blue denotes significant dominance of IDUs and HETs, and dark gray denotes no significant dominance by either transmission group.

pared with MSMs (mean size, 29 individuals; 36% of patients in the 3 largest transmission clusters), suggesting more-independent subepidemics and smaller transmission chains for the MSM epidemic.

The transmission clusters differ considerably with regard to the dates of infection obtained from the back-calculation model. (Figure 2 shows the ranges of dates of infection for the 20 largest clusters.) Transmission clusters were classified as old or young, depending on whether the oldest infection occurred before or after 1990. We found that most patients (96%) in Swiss transmission clusters belong to 55 "old" transmission clusters. However, the phylogeny does also contain 5 "young" transmission clusters, which include 4% of patients that belong to Swiss transmission clusters.

Interactions between transmission groups. Table 2 shows how the transmission groups are distributed over the different cluster types. Overall, 72% of the IDU sequences but only 42% of the MSM sequences belong to Swiss transmission clusters. Thus, MSM sequences clustered particularly poorly, indicating an important role of infections acquired abroad, whereas domestic infection events seem to dominate HIV-1 epidemics among IDUs. Furthermore, Table 2 indicates considerable mixing between the HET and MSM groups: 11% of the HET sequences stem from MSM transmission clusters. By contrast only 2% of IDU sequences were associated with MSM transmission clusters. These findings were further corroborated and extended by results from a multinomial logistic regression model, to investigate associations of demographic factors with IDU/HET transmission clusters or MSM transmission clusters, using unclustered sequences as a reference (Table 3). Clustering of sequences in IDU/HET transmission clusters was associated with a smaller probability of the origin of the patients being outside of Western Europe and a higher probability of patients being female, having injection drug use as a presumed mode of HIV

Table 1. Characterization of the 19 Largest Swiss Transmission Clusters and Nonclustered Individuals in Terms of Size, Recruitment Region, and Transmission Group

Cluster (size)[a]	Subjects with acute infection	Recruitment region[b]		Region of origin		Presumed HIV acquisition group			Female patients	Year of first positive HIV test result,[c] median (IQR)	HIV infection	
		West	East	Other	Western Europe	HET	IDU	MSM			Median year of occurrence (IQR)	Patients, no.
No cluster (n = 2577)	478 (18.5)	780 (30.3)	1797 (69.7)	604 (23.4)	1973 (76.6)	617 (23.9)	490 (19)	1470 (57)	476 (18.5)	1994 (1989 to 2001)	1993 (1988 to 1998)	2051
IDU/HET (n = 1051)	157 (14.9)	241 (22.9)	810 (77.1)	135 (12.8)	916 (87.2)	310 (29.5)	700 (66.6)	41 (3.9)	443 (42.2)	1990 (1986 to 1995)	1990 (1987 to 1994)	859
MSM (n = 263)	36 (13.7)	23 (8.7)	240 (91.3)	36 (13.7)	227 (86.3)	27 (10.3)	10 (3.8)	226 (85.9)	4 (1.5)	1994 (1990 to 1999)	1992 (1988 to 1997)	214
IDU/HET (n = 185)	29 (15.7)	38 (20.5)	147 (79.5)	22 (11.9)	163 (88.1)	64 (34.6)	113 (61.1)	8 (4.3)	76 (41.1)	1989 (1986 to 1994)	1989 (1987 to 1993)	142
IDU/HET (n = 173)	24 (13.9)	51 (29.5)	122 (70.5)	31 (17.9)	142 (82.1)	62 (35.8)	93 (53.8)	18 (10.4)	67 (38.7)	1991 (1988 to 1996)	1990 (1987 to 1994)	132
IDU/HET (n = 165)	37 (22.4)	39 (23.6)	126 (76.4)	23 (13.9)	142 (86.1)	37 (22.4)	114 (69.1)	14 (8.5)	54 (32.7)	1994 (1989 to 1998)	1992 (1988 to 1996)	143
AMBIG (n = 88)	18 (20.5)	12 (13.6)	76 (86.4)	9 (10.2)	79 (89.8)	20 (22.7)	45 (51.1)	23 (26.1)	28 (31.8)	1993 (1987 to 2001)	1991 (1988 to 1996.5)	72
MSM (n = 79)	11 (13.9)	33 (41.8)	46 (58.2)	9 (11.4)	70 (88.6)	14 (17.7)	1 (1.3)	64 (81)	3 (3.8)	1997 (1991 to 2004)	1993 (1988 to 2000)	58
AMBIG (n = 77)	12 (15.6)	19 (24.7)	58 (75.3)	11 (14.3)	66 (85.7)	25 (32.5)	31 (40.3)	21 (27.3)	25 (32.5)	1993 (1988 to 1997)	1990.5 (1987 to 1994)	70
MSM (n = 50)	9 (18)	16 (32)	34 (68)	4 (8)	46 (92)	7 (14)	1 (2)	42 (84)	3 (6)	1995 (19900 to 2001)	1993 (1989 to 1996.5)	40
MSM (n = 47)	5 (10.6)	1 (2.1)	46 (97.9)	13 (28.3)	33 (71.7)	15 (31.9)	1 (2.1)	31 (66)	3 (6.4)	1999 (1994 to 2002)	1995 (1992 to 1998)	42
MSM (n = 39)	14 (35.9)	3 (7.7)	36 (92.3)	6 (15.4)	33 (84.6)	2 (5.1)	0 (0)	37 (94.9)	0 (0)	2004 (2001 to 2006)	2000 (1996 to 2002)	31
MSM (n = 34)	9 (26.5)	6 (17.6)	28 (82.4)	5 (14.7)	29 (85.3)	3 (8.8)	0 (0)	31 (91.2)	0 (0)	2001 (1997 to 2005)	1996.5 (1994 to 2001)	30
MSM (n = 33)	6 (18.2)	2 (6.1)	31 (93.9)	5 (15.2)	28 (84.8)	5 (15.2)	2 (6.1)	26 (78.8)	4 (12.1)	2000 (1992 to 2004)	1998 (1990 to 2001)	28
MSM (n = 32)	4 (12.5)	0 (0)	32 (100)	4 (12.5)	28 (87.5)	7 (21.9)	5 (15.6)	20 (62.5)	4 (12.5)	1995 (1990 to 1997)	1992 (1989 to 1994)	21
IDU/HET (n = 31)	3 (9.7)	12 (38.7)	19 (61.3)	2 (6.5)	29 (93.5)	6 (19.4)	24 (77.4)	1 (3.2)	11 (35.5)	1987 (1985 to 1992)	1988 (1985 to 1991)	21
MSM (n = 30)	5 (16.7)	23 (76.7)	7 (23.3)	5 (16.7)	25 (83.3)	5 (16.7)	2 (6.7)	23 (76.7)	2 (6.7)	1994 (1992 to 1997)	1990 (1988 to 1994)	26
MSM (n = 30)	10 (33.3)	9 (30)	21 (70)	8 (26.7)	22 (73.3)	4 (13.3)	1 (3.3)	25 (83.3)	2 (6.7)	1999 (1995 to 2004)	1997 (1994 to 2002)	25
AMBIG (n = 29)	4 (13.8)	5 (17.2)	24 (82.8)	6 (20.7)	23 (79.3)	10 (34.5)	13 (44.8)	6 (20.7)	6 (20.7)	1992 (1986 to 1996)	1989 (1986 to 1993)	25
MSM (n = 29)	3 (10.3)	0 (0)	29 (100)	8 (27.6)	21 (72.4)	5 (17.2)	2 (6.9)	22 (75.9)	2 (6.9)	1997 (1991 to 2001)	1994 (1991 to 1999)	25

NOTE. Data are the no. (%) of patients, unless otherwise indicated. AMBIG, ambiguous; HET, patients who acquired human immunodeficiency virus (HIV) through heterosexual contact; IDU, injection drug users; IQR, interquartile range; MSM, men who have sex with men.

[a] Number of patients in the transmission clusters.

[b] French- or German-speaking Switzerland.

[c] Information on the year of the first positive HIV test result was available for all patients in the clusters.

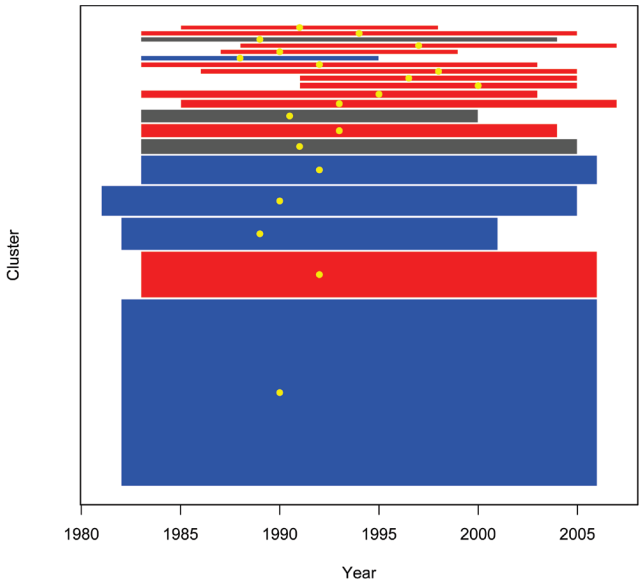

Figure 2. Median *(yellow dots)* value and range *(bars)* of the years of infection for the patients in the 20 largest transmission clusters shown in Table 1. The height of the bars is proportional to the number of patients in the corresponding transmission clusters. Red denotes significant dominance of men who have sex with men (MSM), blue denotes significant dominance of injection drug users (IDUs) and patients acquiring human immunodeficiency virus via heterosexual contact (HETs), and dark gray denotes no significant dominance by either transmission group.

acquisition, and having been infected with HIV before 1996. In contrast, factors associated with MSM transmission clusters were being male, acquisition of HIV through homosexual contact between men, and enrollment into the SHCS in a study center from the German-speaking part of Switzerland.

A particularly interesting aspect of interactions between transmission groups is the source of infection of patients in the HET group. If transmission among HETs would be able to sustain an epidemic in Switzerland independently of other risk groups, we would expect to find large clusters that are dominated by HET sequences (in particular, given that IDUs are, overall, such a small population, compared with HETs). However, the phylogenetic tree contains no such clusters. The largest cluster, with at least 80% HETs, is only 9 sequences large; by contrast, the corresponding clusters for MSMs and IDUs have sizes of 239 sequences and 42 sequences, respectively. (See Figure A1 in Appendix A, which appears only in the electronic version of the *Journal*, for the size distribution of transmission clusters that are dominated by a particular transmission group.) Thus, although HETs are well represented in our database (see Figure A2 in Appendix A, which appears only in the electronic version of the *Journal*), the sequence data show no evidence for large HET-dominated transmission chains in Switzerland. To resolve the origin of the HET sequences, we subdivided them into 3 categories according to their clustering with sequences of other transmission groups (Table 2): (1) those that cluster with IDU sequences (ie, that belong to the IDU/HET transmission clusters), (2) those that cluster with MSM sequences, and (3) those that do not belong to Swiss transmission clusters. IDU- and MSM-associated HET sequences can be interpreted as offshoots of the epidemics in IDUs and MSM, respectively, whereas HET sequences that belong to no larger transmission cluster most plausibly stem from non-Swiss epidemics.

Temporal trends in the interaction between transmission groups. Interestingly, infection dates indicate a shift between the 2 major sources of the HET epidemics (Figure 3). In the early phase of the epidemics, a slightly larger amount of infections (44%) stemmed from the Swiss IDU epidemic than from foreign epidemics (38%). Subsequently, however, the importance of the IDU epidemic decreased considerably such that the association with IDU has almost become negligible, compared with the effect of foreign epidemics (18% versus 71% of sequences belong to the respective categories for 2003–2006). A test for trend shows that the decrease in the effect of the IDU epidemic on the HET epidemic is statistically significant ($P < .001$). Moreover, Figure 3 shows that new HET infections associated with IDUs seem to have decreased in absolute numbers over time, although we cannot fully exclude an effect of delays in diagnosis of HIV infection and enrollment into the SHCS for later time periods. In contrast to HETs, no significant temporal trend was found regarding the origin of MSMs and IDUs (results not shown).

The influence of geographical structure. Concerning the geographical origin of the HIV-1–infected individuals, the strongest segregation can be observed between German- and French-speaking regions. Overall, 74% and 26% of the sequences stem from German- and French-speaking parts of Switzerland, respectively. Patients from German-speaking regions are significantly ($P < .05$ under a binomial null-model) overrepresented in 14 transmission clusters containing 604 patients, whereas patients from French-speaking regions are overrepresented in 10 transmission clusters containing 399 patients. Thus, the phylogeny indicates that geographical struc-

Table 2. Distribution of the Patients from the Different Transmission Groups over the Different Cluster Types

Transmission group	MSM cluster	IDU/HET cluster	Ambiguous cluster	No transmission cluster
MSM	903 (36)	86 (3)	83 (3)	1469 (58)
HET	159 (11)	525 (38)	87 (6)	617 (44)
IDU	37 (2)	1113 (63)	125 (7)	490 (28)

NOTE. Data are the no. (%) of patients. HET, patients acquiring HIV via heterosexual transmission; IDU, injecting drug users; MSM, men having sex with men.

Table 3. Associations of Demographic Factors with Cluster Type, as Inferred from a Multinomial Logistic Regression Model

Factor	No cluster (n = 2577)	IDU/HET cluster (n = 1725)	MSM cluster (n = 1100)	AMBIG cluster (n = 295)	IDU/HET RRR (95% CI)[a]	MSM RRR (95% CI)[a]	AMBIG RRR (95% CI)[a]
ART status at sampling							
Naive	1212 (47.0)	841 (48.8)	553 (50.3)	135 (45.8)	Reference	Reference	Reference
Experienced	1365 (53.0)	884 (51.2)	547 (49.7)	160 (54.2)	0 .76 (.65–0.89)	0 .94 (0.80–1.11)	1 .07 (0.82–1.40)
Years of infection							
1981–1986	354 (13.7)	432 (25.0)	95 (8.6)	53 (18.0)	1 .84 (1.39–2.43)	0 .77 (0.57–1.03)	0 .91 (0.58–1.43)
1987–1991	559 (21.7)	509 (29.5)	191 (17.4)	63 (21.4)	1 .76 (1.37–2.27)	0 .95 (0.75–1.21)	0 .81 (0.54–1.23)
1992–1995	494 (19.2)	320 (18.6)	239 (21.7)	62 (21.0)	1 .55 (1.19–2.00])	1 .17 (0.93–1.48)	1 .01 (0.68–1.51)
1996–2000	483 (18.7)	276 (16.0)	242 (22.0)	53 (18.0)	1 .46 (1.13–1.88)	1 .20 (0.96–1.48)	0 .95 (0.64–1.41)
2001–2007	687 (26.7)	188 (10.9)	333 (30.3)	64 (21.7)	Reference	Reference	Reference
Sampling done during primary HIV infection							
No	2099 (81.5)	1452 (84.2)	907 (82.5)	244 (82.7)	Reference	Reference	Reference
Yes	478 (18.5)	273 (15.8)	193 (17.5)	51 (17.3)	1 .18 (0.96–1.45)	0 .84 (0.69–1.02)	1 .07 (0.77–1.50)
Location of enrollment into SHCS							
French-speaking area	780 (30.3)	403 (23.4)	219 (19.9)	61 (20.7)	0 .83 (0.70–0.98)	0 .59 (0.49–0.70)	0 .67 (0.49–0.90)
German-speaking area	1797 (69.7)	1322 (76.6)	881 (80.1)	234 (79.3)	Reference	Reference	Reference
Region of origin							
Outside Western Europe	604 (23.4)	232 (13.4)	180 (16.4)	47 (15.9)	0 .58 (0.41–0.81)	0 .51 (0.42–0.62)	0 .70 (0.58–0.85)
Western Europe	1973 (76.6)	1493 (86.6)	920 (83.6)	248 (84.1)	Reference	Reference	Reference
Presumed mode of HIV acquisition							
HET	617 (23.9)	526 (30.5)	159 (14.5)	87 (29.5)	Reference	Reference	Reference
IDU	490 (19.0)	1113 (64.5)	38 (3.5)	125 (42.4)	2 .22 (1.87–2.65)	0 .27 (0.18–0.39)	1 .63 (1.18–2.25)
MSM	1470 (57.0)	86 (5.0)	903 (82.1)	83 (28.1)	0 .07 (0.05–.09)	1 .65 (1.31–2.08)	0 .33 (0.23–0.47)
Sex							
Female	476 (18.5)	707 (41.0)	56 (5.1)	77 (26.1)	1 .10 (0.93–1.29)	0 .43 (0.31–0.60)	0 .80 (0.59–1.10)
Male	2101 (81.5)	1018 (59.0)	1044 (94.9)	218 (73.9)	Reference	Reference	Reference

NOTE. Data are the no. (%) of patients, unless otherwise indicated. AMBIG, ambiguous cluster; ART, antiretroviral therapy; CI, confidence interval; HET, patients acquiring HIV via heterosexual transmission; HIV, human immunodeficiency virus; IDU, injecting drug users; MSM, men having sex with men. NA, not applicable; RRR, relative risk ratio; SHCS, Swiss HIV Cohort Study.

[a] 95% confidence intervals not including 1 denote statistically significant associations.

ture, even within a small country such as Switzerland, has a considerable influence on the spread of HIV.

DISCUSSION

Using 5700 subtype B pol sequences from a nationwide and representative database of HIV-1 nucleotide sequences, we characterized the epidemics of HIV-1 in Switzerland since the early 1980s. The broadest pattern observed was the strong separation between the transmission clusters dominated by MSMs and those dominated by IDUs and HETs. Separation of IDUs from other transmission groups (mainly MSMs and hemophiliacs) has previously been found in a number of smaller

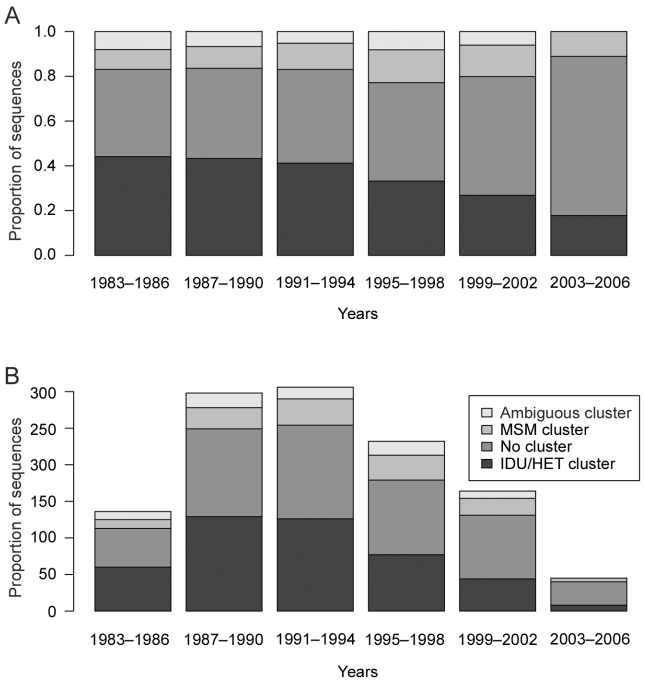

Figure 3. Distribution of sequences from individuals who acquired human immunodeficiency virus via heterosexual contact (HET) over different transmission cluster types as a function of time. *A,* The proportion of HET that belong to a particular transmission cluster type (see key). *B,* The corresponding absolute numbers. IDU, injection drug users; MSM, men who have sex with men,

studies [9, 10, 25, 26], whereas an association between IDUs and HETs has only been documented for a fairly small study based on patients from Edinburgh [9]. On top of this broad separation, we have found a weaker but still substantial exchange between HETs and MSMs, which has not been found in previous molecular epidemiology studies. The transmission clusters also showed a clear signature of geographical separation (in contrast to [25]). However, this signature reached a significant level only in a few transmission clusters, and, overall, geography seemed to have a considerably weaker effect than transmission mode, such as injection drug use or sex between men. In other words, transmission events occur preferentially within the same transmission groups and within the same region, but HIV-1 infections cross geographical regions much more frequently than they cross transmission groups.

In Switzerland, the number of new HIV infections strongly decreased in the late nineties, but increases in incident HIV-1 infections have been noted again in recent years, especially among MSM [27]. Interestingly, we found several transmission clusters that emerged only very recently, with the majority (4 of 5) of those dominated by sequences from MSMs. This finding suggests that the recent increase in the epidemics might, at least in part, be due to recent transmission clusters and not to the growth of the older, established transmission clusters. This finding is in line with recent observations that recently infected patients are an important driver for the HIV epidemic [11, 28–30]. Thus, prevention measures have to target specifically these newer and also younger transmission networks, to be more successful in the future.

Caveats. As in all (molecular) epidemiology studies, sampling could potentially distort our results. For example, it is, in principle, conceivable that the absence of large HET-dominated transmission chains is just an artifact of sampling. This could be the case if our sample systematically omits entire subpopulations of HETs among which large transmission chains have occurred. We think, however, that such a scenario is unlikely for a reportable infection, which (even though after a delay) leads to very severe symptoms. A related issue is that, even under the assumption of a homogenous sampling, we underestimate the length of transmission chains because our sample is incomplete. However, because our sampling density is high for all transmission groups (see Figure A2), this effect will be moderate. Moreover, because the sampling density is comparable for the 3 transmission groups (see Figure A2 in Appendix A, which appears only in the electronic version of the *Journal*), the incompleteness of the sampling does not affect the relative contribution of the different transmission routes. Concerning temporal aspects of the sampling, it should first be noted that the time of infection is only a rough estimate obtained using a back-calculation method (see Materials and Methods). Second, although our data set has an exceptional temporal depth (because the sequences stem not only from drug resistance tests but also from the sequencing of stored plasma samples), the sampling intensity is not homogenous over time and might therefore distort temporal trends of the absolute numbers of patients (such as the pattern shown Figure 3*B*). However, interpretations of relative numbers of patients (such as the key result that the fraction of HETs clustering with IDUs decreases over time) obviously are not affected by these problems. In summary, we think that although sampling might affect the quantitative details of our results, the large qualitative patterns are robust against it. This is not least because of the exceptional completeness and temporal depth of our sample, which is as representative as it gets for epidemics in an entire county over a time span of 2 decades.

We furthermore assessed the robustness of our results both with respect to the method of inferring the phylogenetic tree and with respect to the nucleotide positions considered (see Appendix B, which is available only in the electronic version of the *Journal,* for supplementary material on tests of robustness). These analyses show that the detailed composition of the transmission clusters is not robust; however, on the other hand, they also underline the robustness of the broad qualitative patterns described in this article. Finally, it should be noted that our conclusions are restricted to the HIV-1 subtype B epidemics in Switzerland. This especially concerns our finding that large

transmission chains dominated by HETs are absent in Switzerland. With an increasing number of non–subtype B infections (which occur mainly in migrants), there could be, for instance, a burgeoning epidemic of heterosexually transmitted non–B subtype HIV-1 within the migrant community in Switzerland.

Conclusion. Although the association between IDUs and HETs is, overall, a prominent feature in the Swiss HIV epidemics, our data also show that it has become considerably weaker, especially in comparison to migration. This shift from injection drug use–associated infections to infections from foreign epidemics probably reflects 2 complementary developments. On the one hand, the importance of migration has increased. On the other hand, the epidemics among IDUs have strongly decreased because of effective prevention (eg, needle exchange programs [31, 32]), and, as a secondary effect, fewer HETs have been infected by IDUs. Our analysis indicates that the second factor was most likely responsible for the major part of the shift and thus supports the view that heterosexuals indirectly profited from the prevention among the IDU community. Thus, the decreasing effect of IDUs on the HET epidemic indicates that prevention measures targeted at IDUs, such as needle exchange programs, can result in substantial alleviations for other transmission groups as well.

Acknowledgments

We thank the patients participating in the Swiss HIV Cohort Study for their commitment, all the study nurses and study physicians for their invaluable work, the data center for data management, all the resistance laboratories for their high-quality work, and SmartGene for providing an impeccable database service. We also thank Alexei Drummond and Joseph Wong for critical reading of the manuscript.

Financial support. This study was financed in the framework of the Swiss HIV Cohort Study (SHCS), supported by the Swiss National Science Foundation (SNF grant 3345-062041). Additional support was provided by SNF grants 3247B0-112594 (to H.F.G., S.Y., B.L., and S.B.) and 320000-116035 (to H.F.G.); SHCS projects 470, 528, and 569; the SHCS Research Foundation; and the Union Bank of Switzerland (in the form of an additional research grant in the name of a donor to H.F.G.). The research leading to these results has received funding from the European Community's Seventh Framework Programme (pp7/2007–2013) under the project collaborative HIV and Anti-HIV Drug Resistance Network (CHAIN; grant agreement 223131). V.v.W. was supported by a fellowship from The Novartis Foundation, formerly known as Ciba-Geigy Jubilee Foundation. The funding agencies had no role in conducting the study or in preparing the manuscript.

References

1. Drummond AJ, Pybus OG, Rambaut A, Forsberg R, Rodrigo AG. Measurably evolving populations. Trends Ecol Evol 2003; 18:481–8.
2. Gao F, Bailes E, Robertson DL, et al. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature **1999**; 397:436–41.
3. Gilbert MTP, Rambaut A, Wlasiuk G, et al. The emergence of HIV/AIDS in the Americas and beyond. Proc Natl Acad Sci U S A **2007**; 104:18,566–70.
4. Hué S, Pillay D, Clewley JP, Pybus OG. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. Proc Natl Acad Sci U S A **2005**; 102:4425–9.
5. Korber B, Muldoon M, Theiler J, et al. Timing the ancestor of the HIV-1 pandemic strains. Science **2000**; 288:1789–96.
6. Ou CY, Ciesielski CA, Myers G, et al. Molecular epidemiology of HIV transmission in a dental practice. Science **1992**; 256:1165–71.
7. de Oliveira T, Pybus OG, Rambaut A, et al. Molecular epidemiology-HIV-1 and HCV sequences from Libyan outbreak. Nature **2006**; 444: 836–7.
8. Gifford RJ, de Oliveira T, Rambaut A, et al. Phylogenetic surveillance of viral genetic diversity and the evolving molecular epidemiology of human immunodeficiency virus type 1. J Virol **2007**; 81:13,050–6.
9. Holmes EC, Zhang LQ, Robertson P, et al. The molecular epidemiology of human immunodeficiency virus type-1 in Edinburgh. J Infect Dis **1995**; 171:45–53.
10. Kuiken CL, Cornelissen MTE, Zorgdrager F, Hartman S, Gibbs AJ, Goudsmit J. Consistent risk group-associated differences in human immunodeficiency virus type 1 vpr, vpu and V3 sequences despite independent evolution. J Gen Virol **1996**; 77:783–92.
11. Lewis F, Hughes GJ, Rambaut A, Pozniak A, Brown AJL. Episodic sexual transmission of HIV revealed by molecular phylodynamics. PLoS Med **2008**; 5:e50.
12. Joint United Nations Programme on HIV/AIDS (UNAIDS). AIDS epidemic update (December 2007). UNAIDS, **2007**.
13. Perrin L, Kaiser L, Yerly S. Travel and the spread of HIV-1 genetic variants. Lancet Infect Dis **2003**; 3:22–7.
14. Yerly S, von Wyl V, Ledergerber B, et al. Transmission of HIV-1 drug resistance in Switzerland: a 10-year molecular epidemiology survey. AIDS **2007**; 21:2223–9.
15. Swiss HIV Cohort Study. http://www.shcs.ch. Accessed 19 March **2010**.
16. Ledergerber B, Egger M, Opravil M, et al. Clinical progression and virological failure on highly active antiretroviral therapy in HIV-1 patients: a prospective cohort study. Lancet **1999**; 353:863–8.
17. Von Wyl V, Yerly S, Böni J, et al. Emergence of HIV-1 drug resistance in previously untreated patients initiating combination antiretroviral treatment—a comparison of different regimen types. Arch Intern Med **2007**; 167:1782–90.
18. Yerly S, Vora S, Rizzardi P, et al. Acute HIV infection: impact on the spread of HIV and transmission of drug resistance. AIDS **2001**; 15: 2287–92.
19. Boni J, Pyra H, Gebhardt M, et al. High frequency of non-B subtypes in newly diagnosed HIV-1 infections in Switzerland. J Acquir Immune Defic Syndr **1999**; 22:174–9.
20. Los Alamos HIV Database. http://www.hiv.lanl.gov/content/index. Accessed 19 March **2010**.
21. Hirsch MS, Gunthard HF, Schapiro JM, et al. Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society–USA panel. Clin Infect Dis **2008**; 47:266–85.
22. Taffé P, May M. A joint back calculation model for the imputation of the date of HIV infection in a prevalent cohort. Stat Med **2008**; 27: 4835–53.
23. Stamatakis A. Phylogenetic models of rate heterogeneity: a high performance computing perspective. Rhodos, Greece: **2006**.
24. Stamatakis A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **2006**; 22:2688–90.
25. Brown AJL, Lobidel D, Wade CM, et al. The molecular epidemiology of human immunodeficiency virus type 1 in six cities in Britain and Ireland. Virology **1997**; 235:166–77.
26. Kuiken C, Thakallapalli R, Eskild A, de Ronde A. Genetic analysis reveals epidemiologic patterns in the spread of human immunodeficiency virus. Am J Epidemiol **2000**; 152:814–22.
27. Gebhardt M. Recent trends in new diagnoses of HIV infections in

Switzerland: probable increase in MSM despite an overall decrease. Euro Surveill **2005**; 10: E051208.2.
28. Brenner BG, Roger M, Routy JP, et al. High rates of forward transmission events after acute/early HIV-1 infection. J Infect Dis **2007**; 195: 951–9.
29. Pao D, Fisher M, Hué S, et al. Transmission of HIV-1 during primary infection: relationship to sexual risk and sexually transmitted infections. AIDS **2005**; 19:85–90.
30. Yerly S, Junier T, Gayet-Ageron A, et al. The impact of transmission clusters on primary drug resistance in newly diagnosed HIV-1 infection. AIDS **2009**; 23:1415–23.
31. Nordt C, Stohler R. Incidence of heroin use in Zurich, Switzerland: a treatment case register analysis. Lancet **2006**; 367:1830–4.
32. Rezza G, Rota MC, Buning E, et al. Assessing HIV prevention among injecting drug users in European Community countries: a review. Soz Praventivmed **1994**; 39:S61–78.